

absolutely continuous probability distributions. All the remaining inequalities of the same type either are weaker or coincide with that asymptotically; the Cramér–Rao–Darmois–Fréchet inequality can only be made more precise by higher order corrections.

PROBLEM 6. Find the bound for the remainder term in (23) in terms of the maximum of invariant curvature of the family $\{P_\theta\}$.

Note that under the chosen loss function $2I(Q: P_\theta)$ for exponential families the problem of estimating the law P_θ reduces (due to Theorem 6) to the problem of estimating its parameter, i.e. the estimates of $Q \in \{P_\theta\}$ -type form a complete class.

PROBLEM 7. In what natural terms family smoothness should be described in Theorem 8?

It is completely obscure how to replace the usual sufficient conditions of smoothness in terms of majorant existence in the third derivatives since the analysis then requires to be developed in non-linear, non-topological space.

PROBLEM 8. How does the formulation of Theorem 8 change when the smooth family $\{P_\theta\}$ has points of self-intersection?

Acknowledgement. The author considers it his pleasant duty to thank Professor R. Bartoszyński for his attention.

References

- [1] D. Blackwell, *Ann. Math. Statist.*, 24 (1953), pp. 265–272.
- [2] H. Chernoff, *ibid.* 23 (1952), pp. 493–507.
- [3] H. Cramér, *Mathematical methods of statistics*, Princeton Univ. Press, Princeton 1946.
- [4] I. Csiszar, *Ann. Probab.* 3 (1975), pp. 146–158.
- [5] N. N. Čencov, *Statistical decision rules and optimal inference*, Nauka, Moscow 1972.
- [6] —, *Math. Notes* 4 (1968), pp. 323–332.
- [7] N. P. Salikhov, *DAN USSR* 209 (1972).
- [8] Ch. Stein, *J. Roy. Statist. Soc.* 24 (1962), pp. 265–296.
- [9] A. Wald, *Ann. Math. Statist.* 10 (1939), pp. 299–326.

Presented to the semester
MATHEMATICAL STATISTICS
September 15–December 18, 1976

HSU'S THEOREM IN VARIANCE COMPONENT MODELS*

HILMAR DRYGAS**

J. W. Goethe-Universität, Fachbereich Mathematik
6 Frankfurt am Main, Robert-Mayer-Strasse 6–10, FRG

0. Introduction

This paper deals with linear models of the kind $y = X\theta + U\varepsilon$, where ε is a random vector composed of independent random variables. Therefore $\text{Cov } \varepsilon = \sum_{i=1}^k \sigma_i^2 V_i$, where $V_i = \text{diag}(0, \dots, I_i, \dots, 0)$ and I_i is the unit matrix of appropriate order (or any diagonal matrix). Much work has been done to investigate the problem of existence of uniformly best (invariant) quadratic unbiased estimators if ε is normally or quasi-normally distributed. It is the purpose of this paper to extend these results to the non-normal case. This extension is done in the case of best invariant quadratic unbiased estimators. A complicated matrix-relation turns out to ensure optimality. But in analogy to Hsu's theorem it can be shown that this relation can be replaced by requiring it only for the diagonal elements. The obtained results still appear very complicated but it turns out that due to the diagonality of the V_i the verification of the obtained conditions is rather straightforward. This is illustrated at two examples: the balanced one-way and the balanced two-way classification model.

1. Notation, Hsu's theorem

Let X be an $n \times s$ -matrix, let θ be an $s \times 1$ -vector and U an $n \times r$ -matrix. Consider the linear model

$$(1.1) \quad y = X\theta + U\varepsilon,$$

where $y = (y_1, \dots, y_n)'$, $\varepsilon = (\varepsilon_1, \dots, \varepsilon_r)'$ are random vectors. It is assumed that the components of ε behave up to their moments of order 4 as independent random

* Paper presented at the conference on "Mathematical Statistics" Wisla, Poland, 13.–18.12. 1976.

** OE 3 der GhK, Gesamthochschule Kassel, Department of Mathematics, Heinrich-Plett-Str. 40, D 3500 Kassel, FRG.

variables, i.e.,

$$(1.2) \quad E(\varepsilon_{i_1}^{n_1} \varepsilon_{i_2}^{n_2} \varepsilon_{i_3}^{n_3} \varepsilon_{i_4}^{n_4}) = \prod_{j=1}^4 E \varepsilon_{i_j}^{n_j}$$

for integers $1 \leq i_1 \neq i_2 \neq i_3 \neq i_4 \leq r$ and $0 \leq n_1 + n_2 + n_3 + n_4 \leq 4$. Moreover, we assume that

$$(1.3) \quad E\varepsilon = 0, \quad \text{Cov} \varepsilon = E\varepsilon\varepsilon' = \sum_{i=1}^k \sigma_i^2 V_i,$$

where $V_i = \text{diag}(v_i)$ is a diagonal matrix. In most applications V_i will be of the form

$$(1.4) \quad V_i = \text{diag}(0, \dots, I_i, 0, \dots, 0),$$

where 0 and I_i are zero-matrices and the unit-matrix, respectively, of appropriate order. $\text{diag}(B_1, \dots, B_m)$ is a block-diagonal matrix with blocks B_1, \dots, B_m :

$$(1.5) \quad \text{diag}(B_1, \dots, B_m) = (\delta_{ij} B_i; i, j = 1, 2, \dots, m).$$

The setup for our linear model (1.1) has earlier been studied by several authors, e.g., Kleffe and Pincus [9], [10], Kleffe [8] and C. R. Rao [15]. From our assumptions (1.3), (1.4) we get immediately

$$(1.6) \quad Ey = X\theta, \quad \text{Cov} y = \sum_{i=1}^k \sigma_i^2 UV_i U',$$

$$(1.7) \quad Eyy' = X\theta\theta'X' + \sum_{i=1}^k \sigma_i^2 UV_i U'.$$

The problem in all such situations is, of course, to estimate θ and $\sigma = (\sigma_1^2, \dots, \sigma_k^2)'$ or linear functions of these parameters. While the problem of estimating θ linearly in y in an uniformly good way is rather satisfactorily solved (Drygas [2], Gnot, Klonecki, Zmyslony [5]), the solution concepts for estimating σ quadratically in y are much more complicated. We refer to papers by Seely [17], [18], [19], [20], Seely and Zyskind [21], Zmyslony [22], [23], Gnot, Zmyslony and Klonecki [5], [6], Kleffe and Pincus [9], [10]. The work on MINQUE theory by Rao [13], [14], [16] should also be mentioned in this context.

Most of the authors assume that ε is quasi-normally distributed. It is the purpose of this paper to abandon this assumption. Let

$$(1.8) \quad E(\varepsilon_i^4) = \beta_i \sigma_i^4, \quad \beta_i \geq 1, \quad i = 1, 2, \dots, r.$$

In the quasi-normal case $\beta_i = 3$. Let us introduce the vector

$$(1.9) \quad \beta - 3 \cdot 1_r = (\beta_1 - 3, \dots, \beta_r - 3)'$$

$\beta - 3$ is called the *curtosis* of the random variable ε_i .

We denote by $\text{diag} A$ the diagonal matrix with the same diagonal elements as the matrix A and by $\text{diag} \alpha$ the diagonal matrix whose i th diagonal element is equal to the i th component of the vector α .

Since we want to consider best invariant quadratic unbiased estimators (BIQUE) of functions $f'\sigma$, $f \in R^k$, it is convenient to consider the random matrix

$$(1.10) \quad Z = Myy'M, \quad M = I - P_{\text{Im}(X)} = I - XX^+.$$

Z is a H -valued random element, where H denotes the set of all symmetric $n \times n$ -matrices A meeting the condition $AX = 0$. In H the inner product $\langle A, B \rangle = \text{tr}(AB)$ is used. Indeed, $y' Ay$ is an invariant estimator of $f'\sigma$ iff $AX = 0$ and this is equivalent to $A = MAM$. Thus $y' Ay = y' MAMy = \text{tr}(AMyy'M) = \text{tr}(AZ) = \text{tr}(ZA)$, which shows that $y' Ay$ is a linear function of Z . Now

$$(1.11) \quad EZ = \sum_{i=1}^k \sigma_i^2 MUV_i U'M =: MUVU'M,$$

where

$$(1.12) \quad V = \sum_{i=1}^k \sigma_i^2 V_i,$$

while (see Kleffe and Pincus [8], p. 151) for all $A \in H$:

$$(1.13) \quad \begin{aligned} \text{Var} \text{tr}(ZA) &= \text{Var}(y' MAMy) = \text{Var}(e' MAMe) \\ &= \text{Var}(e' Ae) = 2\text{tr}(UVU' AUVU' A) + \\ &\quad + \text{tr}(UV \text{diag}(U' AU) \text{diag}(\beta - 3 \cdot 1_r) VU' A) \\ &= 2\text{tr}(MUVU' AUVU' MA) + \\ &\quad + \text{tr}(MUV \text{diag}(U' AU) \text{diag}(\beta - 3 \cdot 1_r) VU' MA) \end{aligned}$$

implying

$$(1.14) \quad \begin{aligned} (\text{Cov} Z)A &= 2MUVU' AUVU' M + \\ &\quad + MUV \text{diag}(U' AU) \text{diag}(\beta - 3 \cdot 1_r) VU' M, \quad A \in H. \end{aligned}$$

From this we get immediately:

1.1. THEOREM. (a) $y' Ay$ is BIQUE of $f'\sigma$ iff (i) $A \in H$, (ii) $(\text{tr}(AUV_i U'); i = 1, 2, \dots, k)' = f$, (iii) $2MUVU' AUVU' M + MUV \text{diag}(U' AU) \text{diag}(\beta - 3 \cdot 1_r) VU' M \in \text{span}\{MUV_i U' M; i = 1, 2, \dots, k\} \forall V \in \text{span}\{V_1, \dots, V_k\}$.

(b) If $y' Ay$ is already BIQUE of $f'\sigma$ in the quasi-normal case, i.e., if for all $V \in \text{span}\{V_1, \dots, V_k\}$

(iii') $MUVU' AUVU' M \in \text{span}\{MUV_i U' M; i = 1, 2, \dots, k\}$, then it is BIQUE of $f'\sigma$ iff

(iii'') $MUV \text{diag}(U' AU) \text{diag}(\beta - 3 \cdot 1_r) VU' M \in \text{span}\{MUV_i U' M; i = 1, 2, \dots, k\} \forall V \in \text{span}\{V_1, \dots, V_k\}$.

Proof. Conditions (a) and (b) are evidently the conditions of invariance and unbiasedness, respectively. By the projection theorem (Drygas [1], p. 375) or the theorem of Lehmann-Scheffé, $A \in H$ minimizes

$$(1.15) \quad \text{Var}(y' Ay) = \text{tr}((\text{Cov} Z)A \cdot A)$$

subject to $(\text{tr}(AUV_iU'))$; $i = 1, 2, \dots, k$)' = f if and only if A meets conditions (i), (ii) and

$$(1.16) \quad \text{Cov}((y'Ay), (y'By)) = \text{tr}((\text{Cov}Z)A \cdot B) = 0$$

for all $B \in H$ satisfying $0 = \text{tr}(BUV_iU') = \text{tr}(MUV_iU'MB)$, $i = 1, 2, \dots, k$. Thus by (1.16)

$$(1.17) \quad (\text{Cov}(Z)A) \perp (\text{span}\{MUV_iU'M; i = 1, \dots, k\})^\perp,$$

implying

$$(1.18) \quad (\text{Cov}Z)A \in \text{span}\{MUV_iU'M; i = 1, 2, \dots, k\}$$

and this should hold for all $V \in \text{span}\{V_1, \dots, V_k\}$. This is condition (iii). Assertion (b) of the theorem is now evident. ■

Now condition (iii'') is rather complicated. Hsu in his pioneering paper [7] has shown that in the case $k = 1$ (iii'') is met if it is already met for the diagonal elements of the corresponding matrices, the converse being obviously also true. Up to date proofs of Hsu's theorem using modern tools of mathematical statistics and linear algebra are available in Pukelsheim [12] and Drygas-Hupet [4]. Also in the case of variance components relation (iii'') can be replaced by a condition which is only valid for the diagonal elements.

At first we remark that since M is positive semi-definite (iii'') is equivalent to

$$(1.19) \quad U'MUV\text{diag}(U'AU)\text{diag}(\beta - 3 \cdot 1_r)VU'MU \\ \in \text{span}\{U'MUV_iU'MU; i = 1, \dots, k\} \quad \forall V \in \text{span}\{V_1, \dots, V_k\}.$$

1.2. THEOREM. Let $B = U'MU$, $\text{diag}b = V\text{diag}(U'AU)\text{diag}(\beta - 3 \cdot 1_r)$, $V, A \in H$ and let, moreover, $B * B$ be the Hadamard product of B and B , $B * B = (b_{ij}^2)$; $i, j = 1, \dots, r$. Finally, let v_i be such that $\text{diag}v_i = V_i$ and $b_i = (B * B)v_i$. Then $MUV\text{diag}(U'AU)\text{diag}(\beta - 3 \cdot 1_r)VU'M \in \text{span}\{MUV_iU'M; i = 1, 2, \dots, k\}$ if and only if

$$(1.20) \quad (B * B)b \in \text{span}\{b_1, \dots, b_k\} = \text{span}\{(B * B)v_i, i = 1, \dots, k\}.$$

Proof. We give two proofs of the theorem according to the two approaches available in the case $k = 1$. First, however, we show that condition (1.20) is necessary. By (1.19) condition (iii'') is equivalent to the existence of numbers q_1, \dots, q_k such that

$$(1.21) \quad U'MU\text{diag}bU'MU = \sum_{i=1}^k q_i(U'MU)V_iU'MU$$

or, using $\text{diag}v_i = V_i$, this is equivalent to

$$(1.22) \quad U'MU\text{diag}\left(b - \sum_{i=1}^k q_i v_i\right)U'MU = 0.$$

Clearly this condition implies

$$(1.23) \quad \text{diag}\left(U'MU\text{diag}\left(b - \sum_{i=1}^k q_i v_i\right)U'MU\right) = \text{diag}\left((B * B)\left(b - \sum_{i=1}^k q_i v_i\right)\right) = 0$$

and this is just (1.20).

For the sufficiency of condition (1.20) a first proof will be given which may be called the *P-approach*. P stands for Pukelsheim ([12]) or projection-method. Let now (1.20) be met and consider $C = MU\text{diag}bU'M$, $L = \text{span}\{MUV_iU'M; i = 1, 2, \dots, k\}$. Assume that by (1.20)

$$(1.24) \quad (B * B)b = \sum_{i=1}^k q_i b_i = \sum_{i=1}^k q_i (B * B)v_i.$$

Now our statement is that $P_L C = \sum_{i=1}^k q_i MUV_iU'M$.

In order to prove this we verify, using

$$\text{tr}(\text{diag}A \cdot B) = \text{tr}(\text{diag}A \cdot \text{diag}B) = \text{tr}(A \cdot \text{diag}B),$$

that

$$(1.25) \quad \left\langle C - \sum_{i=1}^k q_i MUV_iU'M, MUV_jU'M \right\rangle = \left\langle C - \sum_{i=1}^k q_i MUV_iU'M, UV_jU' \right\rangle \\ = \text{tr}\left(\left(C - \sum_{i=1}^k q_i MUV_iU'M\right)UV_jU'\right) = \text{tr}\left(U'\left(C - \sum_{i=1}^k q_i MUV_iU'M\right)UV_j\right) \\ = \text{tr}\left(U'MU\left(\text{diag}\left(b - \sum_{i=1}^k q_i v_i\right)U'MU\right)\text{diag}v_j\right) \\ = \text{tr}\left(\text{diag}\left((U'MU)\left(\text{diag}\left(b - \sum_{i=1}^k q_i v_i\right)U'MU\right)\right)\text{diag}v_j\right) \\ = \text{tr}\left(\left(\text{diag}\left((B * B)\left(b - \sum_{i=1}^k q_i v_i\right)\right)\text{diag}v_j\right)\right) = 0.$$

Since this is correct for all j , the relation $P_L C = \sum_{i=1}^k q_i MUV_iU'M$ is verified. Consequently,

$$(1.26) \quad \|C - P_L C\|^2 = \langle (C - P_L C), C \rangle \\ = \text{tr}((C - P_L C)U\text{diag}bU') \\ = \text{tr}(U'(C - P_L C)U\text{diag}b) = \text{tr}(\text{diag}(U'(C - P_L C)U)\text{diag}b).$$

But as above $\text{diag}(U'(C - P_L C)U) = \text{diag}((B * B)(b - \sum_{i=1}^k q_i v_i))$ which vanishes by (1.20). Thus $\|C - P_L C\|^2 = 0$ implying $C = P_L C \in L$. ■

The second proof of the sufficiency of (1.20) is called the *H-approach*. This may stand for Hadamard-matrices or Hupet (see [4]). A class \mathfrak{M} of symmetric matrices is called *annulating* with respect to the diagonal if $B \in \mathfrak{M}$, $\text{diag } B = 0$ implies $B = 0$. An example of such a class is the class of \mathfrak{M} of positive semi-definite (PSD) matrices B . Drygas-Hupet proved in [4] that the class

$$(1.27) \quad \mathfrak{M} = \{C = AAB: A, B \text{ PSD, } A \text{ diagonal}\}$$

is annulating with respect to the diagonal. Since $U'MU$ is PSD, indeed

$$\text{diag} \left((U'MU) \text{diag} \left(b - \sum_{i=1}^k \varrho_i v_i \right) U'MU \right) = \text{diag} \left((U'MU * U'MU) \left(b - \sum_{i=1}^k \varrho_i v_i \right) \right) = 0,$$

i.e., (1.20), implies (1.19) and this finishes the proof of the theorem. ■

Since b depends on V , condition (1.20) has still to be verified for all $V \in \text{span}\{V_1, \dots, V_k\}$. But since V is a linear combination of the V_i it is enough to have (1.20) for all V_i . This leads to condition which is quite analogous to Hsu's original result.

1.4. THEOREM (Hsu's theorem for variance component models). Let c_{ij} be a $r \times 1$ column-vector such that

$$\text{diag } c_{ij} = V_i \text{diag}(U'AU) \text{diag}(\beta - 3 \cdot 1_r) V_j$$

and

$$(1.28) \quad C = (c_{11}, c_{12}, \dots, c_{1k}, \dots, c_{k1}, \dots, c_{kk}), \quad D = (v_1, \dots, v_k).$$

Then necessary and sufficient for $y'Ay$ to be *BIQUE* of $f'\sigma$, provided $y'Ay$ is already *BIQUE* of $f'\sigma$ in the quasi-normal case, is the existence of a $k \times k^2$ -matrix P such that

$$(1.29) \quad (U'MU * U'MU)(C - DP) = 0.$$

1.5. Remark. In almost all applications $V_i V_j = \delta_{ij} V_i^2$ and therefore $c_{ij} = 0$ if $i \neq j$. In this case C can be replaced by $(c_1 \dots c_k)$ and P is a $k \times k^2$ -matrix.

Proof of Theorem 1.4. Since b is defined via

$$\text{diag } b = V \text{diag}(U'AU) \text{diag}(\beta - 3 \cdot 1_r) V \quad \text{and} \quad V = \sum_{i=1}^k \sigma_i^2 V_i,$$

we get

$$\text{diag } b = \sum_{i,j=1}^k \sigma_i^2 \sigma_j^2 V_i \text{diag}(U'AU) \text{diag}(\beta - 3 \cdot 1_r) V_j = \sum_{i,j=1}^k \sigma_i^2 \sigma_j^2 \text{diag}(c_{ij}),$$

$$b = \sum_{i,j=1}^k \sigma_i^2 \sigma_j^2 c_{ij}.$$

Thus (1.20) is equivalent to

$$(1.30) \quad (B * B)c_l \in \text{span}\{(B * B)v_i, i = 1, \dots, k\}, \quad l = 1, 2, \dots, k^2.$$

(The index (i, j) has been replaced by the index $l = k(i-1) + j$.) This means that

$$(1.31) \quad (B * B)c_l = \sum_{j=1}^k \varrho_{jl}(B * B)v_j, \quad l = 1, 2, \dots, k^2.$$

Let $P = (\varrho_{jl}; j = 1, \dots, k; l = 1, 2, \dots, k^2)$. Then evidently (1.31) is equivalent to (1.29). ■

1.5. THEOREM. Let the model $y = X\theta + U\varepsilon$ be given and assume that $I_n \in \text{span}\{UV_i U'; i = 1, 2, \dots, k\}$. Suppose, moreover, that the quadratic subspace condition for $\text{span}\{MUV_i U' M; i = 1, 2, \dots, k\}$ is satisfied and therefore a *BIQUE* of $f'\sigma$ for any estimable $f'\sigma$ exists in the quasi-normal case. Then a *BIQUE* for any estimable function $f'\sigma$ exists in the non-normal case iff (1.29) is met for all $A \in \text{span}\{MUV_i U' M\}$.

Proof. In Drygas [2] it has been shown that a *BIQUE* of $f'\sigma$ exists for all estimable $f'\sigma$ if and only if a Gauss-Markov estimator (GME) of EZ in the model

$$(1.32) \quad EZ = \sum_{i=1}^k \sigma_i^2 MUV_i U' M,$$

$$(\text{Cov } Z)A = 2(MUVU'AUUVU'M) + MUV \text{diag}(U'AU) \text{diag}(\beta - 3 \cdot 1_r) VU M$$

exists. Since a GME exists if $\beta = 3 \cdot 1_r$ (implying the above mentioned quadratic subspace condition) this GME is the least squares estimator in view of $I_n \in \text{span}\{UV_i U'; i = 1, 2, \dots, n\}$. This estimator is obtained by solving the "normal equations"

$$(1.33) \quad \sum_{j=1}^k \hat{\sigma}_j^2 \text{tr}(MUV_i U' UV_j U') = \text{tr}(UV_i U' Z); \quad i = 1, 2, \dots, k.$$

By Kruskal's theorem ([11]) the least squares estimator is GME if and only if the linear space of expectations of Z , i.e., $L = \text{span}\{MUV_i U' M; i = 1, 2, \dots, k\}$ is left invariant by the covariance-operator. Since the first term of the covariance-operator leaves L invariant due to the existence of a GME in the quasi-normal case, a GME in the non-normal case exists iff for all $V \in \text{span}\{V_1, \dots, V_k\}$:

$$(1.34) \quad MUV \text{diag}(U'AU) \text{diag}(\beta - 3 \cdot 1_r) VU' M \in L \quad \forall A \in L.$$

By Hsu's Theorem 1.4 this is equivalent to (1.28), (1.29). ■

2. Applications, examples

(A) Balanced one-way classification

Let $y_{ij} = \mu + a_i + \varepsilon_{ij}$; $j = 1, 2, \dots, n$; $i = 1, 2, \dots, m$. a_i and ε_{ij} are independent random variables with expectation zero. μ is a constant to be estimated. It is assumed that

$$(2.1) \quad E(a_i^2) = \sigma_a^2; \quad i = 1, 2, \dots, m.$$

$$(2.2) \quad E(\varepsilon_{ij}^2) = \sigma_\varepsilon^2; \quad j = 1, 2, \dots, n; \quad i = 1, 2, \dots, m.$$

We denote by $A \circ B$ the Kronecker product of the matrices A and B , $A \circ B = (a_{ij}B)$. The rule $(A \circ B)(C \circ D) = (AC \circ BD)$ will be used several times in the sequel. Let, moreover, denote by I_t the unit matrix of order t and by 1_t the $t \times 1$ -vector whose entries are all equal to one. With this notation

$$(2.3) \quad y = (y_{11}, \dots, y_{1n}, y_{21}, \dots, y_{2n}, y_{m1}, \dots, y_{mn})' = X\theta + \varepsilon,$$

where

$$(2.4) \quad \theta = \mu, \quad X = 1_m \circ 1_n = 1_{nm},$$

$$(2.5) \quad U = (I_m \circ 1_n \quad I_m \circ I_n),$$

$$(2.6) \quad \varepsilon = (a_1, \dots, a_m; \varepsilon_{11}, \dots, \varepsilon_{1n}, \dots, \varepsilon_{m1}, \dots, \varepsilon_{mn})'.$$

Then U is of order $m \times m + nm =: m \times r$ and

$$(2.7) \quad V_1 = \text{diag}(I_m, 0_m \circ I_n), \quad V_2 = (0_m, I_m \circ I_n),$$

$$(2.8) \quad UV_1 U' = I_m \circ 1_n 1_n', \quad UV_2 U' = I_{nm}; \quad V_1 V_2 = 0.$$

This shows that $I_{nm} \in \text{span}\{UV_1 U', UV_2 U'\}$. Let us introduce $P_n = n^{-1}1_n 1_n'$, $M_n = I_n - P_n = I_n - n^{-1}1_n 1_n'$ and, similarly, $P_m = m^{-1}1_m 1_m'$, $M_m = I_m - m^{-1}1_m 1_m'$. P_n and P_m are the projections onto $\text{im}(1_n)$ and $\text{im}(1_m)$, respectively. Evidently,

$$(2.9) \quad M = I_m \circ I_n - P_m \circ P_n = I_{nm} - (mn)^{-1}1_m 1_m' \circ 1_n 1_n',$$

$$(2.10) \quad MUV_1 U' M = MUV_1 U' = nM_m \circ P_n,$$

$$(2.11) \quad MUV_2 U' M = M,$$

$$(2.12) \quad U' M U = \begin{bmatrix} nM_m & M_m \circ 1_n \\ M_m \circ 1_n & M \end{bmatrix},$$

$$(2.13) \quad U' M U * U' M U = \begin{bmatrix} n^2(M_m * M_m) & (M_m * M_m) \circ 1_n' \\ (M_m * M_m) \circ 1_n & M * M \end{bmatrix},$$

$$(2.14) \quad U'(I_m \circ M_n)U = \begin{bmatrix} 0 & 0 \\ 0 & I_m \circ M_n \end{bmatrix},$$

$$(2.15) \quad U'(M_m \circ P_n)U = \begin{bmatrix} nM_m & M_m \circ 1_n' \\ M_m \circ 1_n & M_m \circ P_n \end{bmatrix}.$$

Let us now split up $\beta - 3 \cdot 1_r$ as follows:

$$(2.16) \quad \beta - 3 \cdot 1_r = ((\beta_1 - 3 \cdot 1_m)', (\beta_2 - 3 \cdot 1_{nm})')'.$$

Then for $A = I_m \circ M_n$

$$(2.17) \quad c_1 = 0, \quad c_2 = (0; (n-1)n^{-1}(\beta_2 - 3 \cdot 1_{nm})')', \quad A = M_m \circ P_n$$

yields,

$$(2.18) \quad c_1 = (n(m-1)m^{-1}(\beta_1 - 3 \cdot 1_m)', 0')',$$

$$c_2 = (0', (nm)^{-1}(m-1)(\beta_2 - 3 \cdot 1_{nm})')'$$

and finally for $A = M$ we get:

$$(2.19) \quad \begin{aligned} c_1 &= (n(m-1)m^{-1}(\beta_1 - 3 \cdot 1_m)', 0')', \\ c_2 &= (0', (nm-1)(nm)^{-1}(\beta_2 - 3 \cdot 1_{nm})')'. \end{aligned}$$

Since $I_m \circ M_n, M, M_m \circ P_n = 0$ if $nm = 1$, i.e., $n = m = 1$, $nm > 1$ can be assumed. Similarly when considering $M_m \circ P_n$ and $I_m \circ M_n$, $m > 1$ and $n > 1$, respectively, can be assumed. Now note that $M_t * M_t = t^{-2}((t-1)^2 - 1)I_t + t^{-2}1_t 1_t'$ is regular if $t > 2$. If $t = 1$ $(M_t * M_t)\beta = 2^{-2}1_2 1_2' \beta = 2^{-2}(1_2' \beta)1_2 \in \text{span}\{1_2\}$. Thus we get

2.1. THEOREM. Consider the model

$$(2.20) \quad Ey = \mu 1_{nm}, \quad \text{Cov} y = \sigma_1^2 I_m \circ 1_n 1_n' + \sigma_2^2 I_m \circ I_n,$$

$$(a) \quad y'(I_m \circ M_n) y \text{ is BIQUE of the expectation } \sigma_2^2 m(n-1)n^{-1}$$

iff

$$(2.21) \quad \beta_2 = \gamma_0 1_{nm} \quad \text{for some } \gamma_0 \in R$$

provided $nm > 2$. If $nm = 1, 2$, then β_2 is allowed to be completely arbitrary.

(b) There exist BIQUE of all estimable functions of σ_1^2, σ_2^2 iff

$$(2.22) \quad \beta_1 = \gamma_1 1_m, \quad \beta_2 = \gamma_2 1_{nm}$$

for some constants γ_1, γ_2 provided $nm > 2$ and $m > 2$.

If $m \leq 2$, then β_1 may be completely arbitrary. If $nm \leq 2$, then β_2 may be completely arbitrary.

Proof. (a) We split up $\beta_2 - 3 \cdot 1_{nm} = ((\beta_{21} - 3 \cdot 1_m)', \dots, (\beta_{2n} - 3 \cdot 1_m)')'$.

Then $c_2 = \delta_0(0', (\beta_2 - 3 \cdot 1_{nm})')'$ and we consider the Hsu condition

$$(2.23) \quad 0 = (U' M U * U' M U)(c_2 - \varrho_1 v_1 - \varrho_2 v_2) \\ = \begin{bmatrix} (M_m * M_m) \left(\alpha_0 \sum_{i=1}^n (\beta_{1i} - 3 \cdot 1_m) - \varrho_1 n^2 1_m - \varrho_2 n 1_m \right) \\ (M * M) (\alpha_0 (\beta_2 - 3 \cdot 1_{nm}) - \varrho_1' 1_{nm} - \varrho_2 1_{nm}) \end{bmatrix},$$

where $\varrho_1'(nm-1)(nm)^{-1} = \varrho_1(m-1)m^{-1}$. (The case $nm = 1$ is trivial.) If $nm = 1$ or 2 this expression vanishes automatically while $(M * M)$ is regular if $nm > 2$. Since then $\alpha_0 \neq 0$ indeed the second line of (2.23) implies $\beta_2 = \gamma_0 1_{nm}$, $\beta_{2i} = \gamma_0 1_m$ for some $\gamma_0 \in R$. But then the first line is met automatically with $\varrho_1 = 0$, $\varrho_2 = \alpha_0(\gamma_0 - 3)$.

(b) The cases $A = M_m \circ P_n$ and $A = M$ parallel completely. In both cases c_i is of the form $c_i = \alpha_i((\beta_1 - 3 \cdot 1_m)', 0')'$, $c_2 = \alpha_2(0', (\beta_2 - 3 \cdot 1_{nm})')'$. The case of c_2 parallels part (a) of the proof and will therefore yield the necessary and sufficient condition $\beta_2 = \gamma_2 1_{nm}$ for some $\gamma_2 \in R$. For the case of c_1 we get the Hsu condition

$$(2.24) \quad 0 = (U' M U * U' M U)(c_1 - \varrho_1 v_1 - \varrho_2 v_2) \\ = \begin{bmatrix} n^2(M_m * M_m) (\alpha_1 (\beta_1 - 3 \cdot 1_m) - \varrho_1 1_m - \varrho_2 n^{-1} 1_m) \\ (M_m * M_m) (\alpha_1 (\beta_1 - 3 \cdot 1_m) - \varrho_1 1_m - \varrho_2' 1_m) \circ 1_n \end{bmatrix},$$

where q'_2 is a suitable multiple of q_2 . If $m = 1, 2$, β_1 may be completely arbitrary while $(M_m * M_m)$ is regular implying $\beta_1 = \gamma_1 1_m$. (2.24) is then met with $q_1 = \alpha_1(\gamma_1 - 3)$, $q_2 = 0$. This finishes the proof of the theorem, since it is known that the quadratic subspace condition is fulfilled for this model, Q.E.D.

Theorem 2.1 extends in some sense results obtained by Kleffe and Pincus [9].

(B) *Balanced two-way classification*

Let $y_{ij} = \mu + a_i + b_j + \varepsilon_{ij}$; $i = 1, 2, \dots, m$; $j = 1, 2, \dots, n$. Here $a_1, \dots, a_m, b_1, \dots, b_n, \varepsilon_{11}, \dots, \varepsilon_{mn}$ are again independent random variables with expectation zero, μ an unknown constant. This example has also been studied in some detail by Kleffe [8]. Let again $y = (y_{11}, \dots, y_{1n}, \dots, y_{m1}, \dots, y_{mn})'$, $X = 1_{mn}$, $\theta = \mu$, $\varepsilon = (a_1, \dots, a_m; b_1, \dots, b_n; \varepsilon_{11}, \dots, \varepsilon_{mn})'$. Then

$$(2.25) \quad y = X\theta + U\varepsilon,$$

where

$$(2.26) \quad U = (I_m \circ 1_n; 1_m \circ I_n; I_m \circ I_n).$$

Here, again, $M = I - (nm)^{-1} 1_m 1'_m \circ 1_n 1'_n$ and

$$(2.27) \quad V_1 = \text{diag}(I_m, 0, 0), \quad V_2 = \text{diag}(0, I_n, 0), \quad V_3 = \text{diag}(0, 0, I_m \circ I_n).$$

Again M and $P = I - M$ commutes with UV_1U' , UV_2U' , $I_{mn} = UV_3U'$ and there-

fore $Py = \bar{y}_{00} = \frac{1}{nm} \sum_{i=1}^m \sum_{j=1}^n y_{ij}$ is BLUE of Ey . We now get that $I_{mn} \in \text{span} \{UV_1U', i = 1, 2, 3\}$ and

$$(2.28) \quad \begin{aligned} MUV_1U'M &= MUV_1U' = n(M_m \circ P_n), \\ MUV_2U'M &= MUV_2U' = m(P_m \circ M_n), \quad MUV_3U'M = M. \end{aligned}$$

Kleffe [8] has shown that $L = \text{span}\{P_m \circ M_n, M_m \circ P_n, M\}$ is a quadratic subspace. Therefore in the quasi-normal case for any $A \in L$, $y'Ay$ is BIQUE of its expectation. Again the question arises whether $y'Ay$ is BIQUE of its expectation in the non-normal case. We are going to investigate this question for $A = M_m \circ P_n, P_m \circ M_n, M$ and $M_m \circ M_n$. Straightforward calculations yield

$$(2.29) \quad U'MU = \begin{bmatrix} nM_m & 0 & M_m \circ 1'_n \\ 0 & mM_n & 1'_m \circ M_n \\ M_n \circ 1_n & 1_m \circ M_n & M \end{bmatrix},$$

$$(2.30) \quad U'MU * U'MU = \begin{bmatrix} n^2 M_m * M_m & 0 & (M_m * M_m) \circ 1'_n \\ 0 & m^2 M_n * M_n & 1'_m \circ M_n * M_n \\ (M_m * M_m) \circ 1_n & 1_m \circ M_n * M_n & M * M \end{bmatrix}.$$

We now again split up $\beta - 3 \cdot 1$, as follows:

$$(2.31) \quad \beta - 3 \cdot 1_r = ((\beta_1 - 3 \cdot 1_m)', (\beta_2 - 3 \cdot 1_n)', (\beta_3 - 3 \cdot 1_{nm})')'.$$

Then we get for $A = M_m \circ M_n$, since $V_i V_j = \delta_{ij} V_i$:

$$(2.31) \quad U'AU = \text{diag}(0, 0, M_m \circ M_n),$$

$$(2.32) \quad \begin{aligned} c_1 &= 0, \quad c_2 = 0, \quad c_3 = (0', 0', (n-1)(m-1)(nm)^{-1}(\beta_3 - 3 \cdot 1_{nm})')' \\ &=: \alpha_0(0', 0', (\beta_3 - 1_{nm})')'. \end{aligned}$$

If we again split up $(\beta_3 - 3 \cdot 1_{nm})' = ((\beta_{3i} - 3 \cdot 1_m)', i = 1, \dots, n)'$, then the Hsu condition is obtained as

$$(2.33) \quad \begin{aligned} 0 &= (U'MU * U'MU)(c_3 - q_1 v_1 - q_2 v_2 - q_3 v_3) \\ &= \left[\left[(M_m * M_m) \left(\alpha_0 \sum_{i=1}^n (\beta_{3i} - 3 \cdot 1_m) - n^2 q_1 1_m - n q_3 1_n \right) \right]' \right. \\ &\quad \left. \left[(M_n * M_n) \left(\alpha_0 \sum_{j=1}^m (\beta_3^{(j)} - 3 \cdot 1_n) - m^2 q_2 1_n - m q_3 1_n \right) \right]' \right. \\ &\quad \left. [M * M(\alpha_0(\beta_3 - 3 \cdot 1_{nm}) - q'_1 1_{nm} - q'_2 1_{nm} - q_3 1_{nm})]' \right]', \end{aligned}$$

$\beta_3^{(j)}$, q'_1 and q'_2 have to be chosen appropriately. If $nm = 1$ or 2 this expression vanishes automatically while $M * M$ is regular if $nm > 2$. Since $\alpha_0 \neq 0$ indeed the third line of (2.33) implies $\beta_3 = \gamma_3 1_{nm}$, $\beta_{3i} = \gamma_3 1_m$, $\beta_3^{(j)} = \gamma_3 1_n$ for some $\gamma_3 \in R$. But then the first two relations of (2.33) are met automatically.

Let us now turn to $A = M_m \circ P_n$. Then

$$(2.34) \quad \begin{aligned} \text{diag}(U'AU) &= \text{diag}(n \text{diag}(M_m), 0, \text{diag}(M_m \circ P_n)) \\ &= \text{diag}(n(m-1)m^{-1}1'_m, 0', (m-1)(nm)^{-1}1'_{nm}), \\ (2.35)_a \quad c_1 &= (n(m-1)m^{-1}(\beta_1 - 3 \cdot 1_m)', 0', 0')', \\ c_2 &= 0, \quad c_3 = (0', 0', (m-1)(nm)^{-1}1'_{nm}). \end{aligned}$$

Similar results are obtained for $A = P_m \circ M_n$, namely

$$(2.36) \quad \text{diag}U'AU = \text{diag}(0', m(n-1)m^{-1}1'_n, (n-1)(nm)^{-1}1'_{nm}),$$

$$(2.37) \quad \begin{aligned} c_1 &= 0, \quad c_2 = (0', m(n-1)n^{-1}(\beta_2 - 3 \cdot 1_n)', 0')', \\ c_3 &= (0', 0', (n-1)(nm)^{-1}1'_{nm}). \end{aligned}$$

Finally $A = M$ yields

$$(2.38) \quad \begin{aligned} c_1 &= (n(m-1)m^{-1}(\beta_1 - 3 \cdot 1_m)', 0', 0')', \\ c_2 &= (0', m(n-1)n^{-1}(\beta_2 - 3 \cdot 1_n)', 0')', \\ c_3 &= (0', 0', (nm-1)(nm)^{-1}(\beta_3 - 3 \cdot 1_{nm})'). \end{aligned}$$

c_3 is always of the form $\alpha_0(0', 0', (\beta_3 - 3 \cdot 1_{nm})')'$. This case has already been discussed above. It remains to study the cases

$$c_1 = \alpha_0((\beta_1 - 3 \cdot 1_m)', 0', 0')' \quad \text{and} \quad c_2 = \alpha_0(0', (\beta_2 - 3 \cdot 1_n)', 0')'.$$

In the first case the Hsu condition is

$$(2.39) \quad \begin{aligned} 0 &= (U'MU * U'MU)(c_1 - p_1 v_1 - p_2 v_2 - q_3 v_3) \\ &= \left[\left[(M_m * M_m)(n^2 \alpha_0(\beta_1 - 3 \cdot 1_m) - n^2 q_1 1_m - q_3 n 1_m) \right]' \right. \\ &\quad \left. [(M_n * M_n)(-m^2 q_2 1_n - m q_3 1_n)]' \right. \\ &\quad \left. ((M_m * M_m(\alpha_0(\beta_1 - 3 \cdot 1_m) - q_1 1_m - q'_2 1_n - q_3 1_{nm}) \circ 1_n))' \right]', \end{aligned}$$

while in the second case we get

$$(2.40) \quad 0 = (U'MU \circ U'MU)(c_2 - \varrho_1 v_1 - \varrho_2 v_2 - \varrho_3 v_3) \\ = [(M_m \circ M_m(-\varrho_1 n^2 1_m - \varrho_3 n 1_m))', \\ [M_n \circ M_n(\alpha_0(\beta_2 - 3 \cdot 1_n) - \varrho_2 n^2 1_n - n\varrho_3 1_n)]', \\ (1_m \circ M_n \circ M_n(\varrho_0(\beta_2 - 3 \cdot 1_n) - \varrho_1 1_n - \varrho_2 1_n - \varrho_3 1_n))']'.$$

If $m \leq 2$, in the first case (2.39) is met automatically if we let $\varrho_2 = \varrho_3 = \varrho'_2 = \varrho'_3 = 0$ and suitable ϱ_1 . If $m > 2$, then $M_m \circ M_m$ is regular and (2.39) is satisfied iff $\beta_1 = \gamma_1 1_m$ for some $\gamma_1 \in R$, $\varrho_1 = \alpha_0(\gamma_1 - 3)$, $\varrho_2 = \varrho'_2 = \varrho_3 = \varrho'_3 = 0$ being a possible choice. Similarly we get the necessary and sufficient condition $\beta_2 = \gamma_2 1_n$ in the second case provided $n > 2$.

2.2. THEOREM. Consider the model $Ey = \mu 1_{nm}$,

$$\text{Cov } y = \sigma_1^2 I_m \circ 1_n 1'_n + \sigma_2^2 1_m 1'_m \circ I_n + \sigma_3^2 I_m \circ I_n.$$

(a) $y'(M_m \circ M_n)y$ is BIQUE of its expectation $\sigma_3^2(n-1)(m-1)(nm)^{-1}$ iff

$$(2.41) \quad \beta_3 = \gamma_3 1_{nm} \quad \text{for some } \gamma_3 \in R,$$

provided $nm > 2$. If $nm \leq 2$, β_3 may be completely arbitrary.

(b) There exists a BIQUE of all estimable functions of σ_1^2 , σ_2^2 , σ_3^2 iff

$$(2.42) \quad \beta_1 = \gamma_1 1_m, \quad \beta_2 = \gamma_2 1_n, \quad \beta_3 = \gamma_3 1_{nm}$$

for some constants $\gamma_1, \gamma_2, \gamma_3$ provided nm, m and n are all greater than two. If $nm \leq 2$, then β_1 may be completely arbitrary. If $m \leq 2$, then β_2 may be completely arbitrary.

3. Extensions

This paper deals with best invariant quadratic unbiased estimation in variance component models. Kleffe and Pincus [9] and Kleffe [8] have already discussed the problem of best quadratic unbiased estimation in the one-way and the two-way balanced classification model. Drygas and Średnicka [3] have extended Hsu's result to the case of best quadratic unbiased estimation in the case of one variance-component. The problem of a characterization for best quadratic unbiased estimators in non-normal variance component models is under investigation and will be dealt with in a subsequent paper.

References

- [1] H. Drygas, *The estimation of residual variance in regression analysis*, Math. Operationsf. Statist. 3 (1972), pp. 373-388.
- [2] —, *Best quadratic unbiased estimation in variance-covariance component models*, Paper presented at the 2nd International Summer School on Model Choice and Parameter Estimation in Regression Analysis, Friedrichsroda-Reinhardtsbrunn (Thuringia/DDR), 8.-18.11.1975, ibid., 8(1977); pp. 211-231.

- [3] —, and M. Średnicka, *A new result in Hsu's model of regression analysis*, Bull. Polish. Acad. Sci. 24 (1976), pp. 1133-1136.
- [4] —, and G. Hupe, *A new proof of Hsu's theorem in regression analysis — A coordinate-free approach*, Math. Operationsf. Statist., 8 (1977), pp. 333-335.
- [5] S. Gnot, W. Klonecki, and R. Zmysłony, *Uniformly minimum variance unbiased estimation in Euclidean spaces*, Paper presented at the 2nd International Summer on Problems of Model Choice and Parameter Estimation in Regression Analysis, Friedrichsroda-Reinhardtsbrunn (Thuringia, DDR), 8.-18.11.1975, ibid., 8 (1977), pp. 199-210.
- [6] —, —, —, *Best linear plus quadratic unbiased estimation for Mixed Linear Models*, Zastos. Mat. 15 (1977), pp. 455-461.
- [7] P. L. Hsu, *On the best quadratic estimate of variance*, Statist. Res. Mem. 2, (1938), pp. 91-104.
- [8] J. Kleffe, *Best Quadratic Unbiased Estimators for Variance Components in the balanced Two-way Classification Model*, Math. Operationsf. Statist. 6 (1975), pp. 189-196.
- [9] —, and R. Pincus, *Bayes and Best Quadratic Unbiased Estimators for Variance component and Heteroscedastic Variances in Linear Models*, ibid. 5 (1974), pp. 147-159.
- [10] —, —, *Bayes and Best Quadratic Unbiased Estimators for Parameters of the Covariance Matrix in a Normal Linear Model*, ibid. 5 (1974), pp. 43-67.
- [11] W. Kruskal, *When are Least Squares and Gauss-Markov Estimators Identical? — A coordinate-free approach*, Ann. Math. Statist. 39 (1968), pp. 70-75.
- [12] F. Pukelsheim, *On Hsu's Model in Regression Analysis*, Math. Operationsf. Statist., 8 (1977), pp. 323-331.
- [13] C. R. Rao, *Estimation of Heteroscedastic Variances in linear Models*, Journ. Amer. Statist. Assoc. 65 (1970), pp. 161-172.
- [14] —, *Estimation of Variance and Covariance Components — MINQUE Theory*, Journ. of Multiv. Analysis 1 (1971), pp. 257-275.
- [15] —, *Minimum variance Quadratic Unbiased Estimation of Variance Components*, ibid. 1 (1971), pp. 445-456.
- [16] —, *Estimation of Variance and Covariance Components in Linear Models*, Journ. Amer. Statist. Assoc. 67 (1972), pp. 112-115.
- [17] J. Seely, *Linear spaces and unbiased estimation*, Ann. Math. Statist. 41 (1970), pp. 1725-1734.
- [18] —, *Linear spaces and unbiased estimation application to the mixed linear model*, ibid. 41 (1970), pp. 1735-1745.
- [19] —, *Quadratic subspaces and completeness*, ibidem 42 (1971), pp. 710-721.
- [20] —, *Completeness for a Family of Multivariate Normal Distributions*, ibid. 43 (1972), pp. 1644-1647.
- [21] —, and G. Zyskind, *Linear spaces and minimum variance unbiased estimation*, ibid. 42 (1971), pp. 691-703.
- [22] R. Zmysłony, *Estimation of variance components in random models*, Zastos. Mat. 13 (1973), pp. 521-527.
- [23] —, *On estimation of parameters in linear models*, ibid. 15 (1977), pp. 271-276.

Presented to the semester
MATHEMATICAL STATISTICS
September 15-December 18, 1976